PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : C12N 15/00

A1

(11) International Publication Number: WO 87/ 00861

(43) International Publication Date: 12 February 1987 (12.02.87)

(21) International Application Number: PCT/EP86/00451

(22) International Filing Date: 29 July 1986 (29.07.86)

(31) Priority Application Number: 85810354.2 (EP)

(32) Priority Date: 31 July 1985 (31.07.85)

(33) Priority Countries: CH, et al.

(71) Applicant *(for all designated States except US)*: BATTELLE MEMORIAL INSTITUTE [US/CH]; 7, route de Drize, CH-1227 Carouge (CH).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)* : BROMLEY, Peter [GB/CH]; 26, chemin du Pont de Ville, CH-1224 Chêne-Bougeries (CH). DREANO, Michel [FR/CH]; 5, chemin des Bouvreuils, CH-1234 Vessy (CH). VOELLMY, Richard [CH/US]; 7240 SW, 124th Street, Miami, FL 33156 (US).

(74) Agents: DOUSSE, Blasco et al.; 7, route de Drize, CH-1227 Carouge (CH).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), DK, FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), NO, SE (European patent), US.

Published

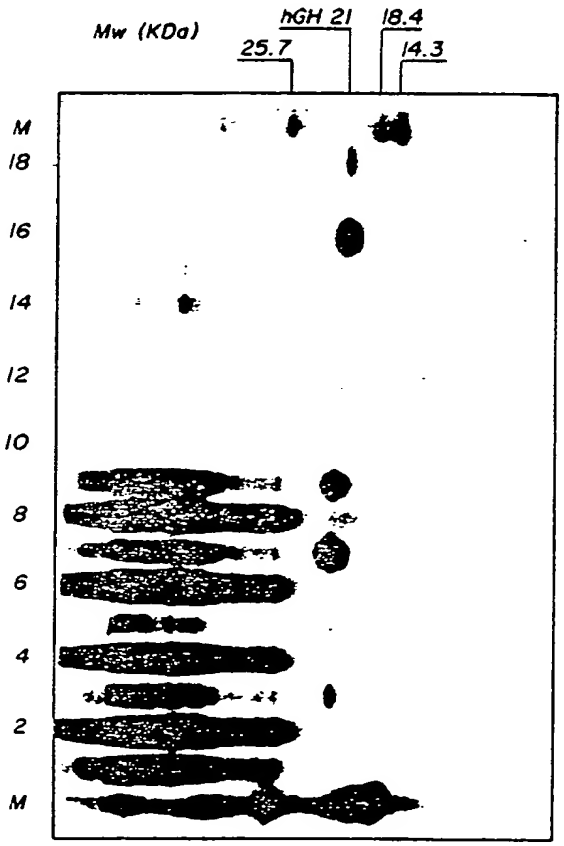
With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: AN IMPROVED HEAT-SHOCK CONTROL METHOD AND SYSTEM FOR THE PRODUCTION OF COMPETENT EUKARYOTIC GENE PRODUCTS

(57) Abstract

In the 70 KDa hsp promoter, modifications brought about to the 5′ and 3′ non-translated sequences, including the insertion or deletion of nucleotides and the number of heat shock consensus sequence elements present in the proximal region of a heat shock gene promoter critically influences promoter activity. Sequence variants in translation initiation and 3′-non-translated regions have been shown to improve levels of induced protein production. Mutant promoters were constructed that have higher maximal activities than the natural parent promoter in heat-shocked cells. These mutant promoters are also expressed at high levels at a temperature below that required for optimal induction of the natural heat-shock genes. These results are of practical interest for attempts to use heat-shock promoters for the inducible production of eukaryotic proteins of medical or industrial importance.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GA**	Gabon	**MR**	Mauritania
AU	Australia	**GB**	United Kingdom	**MW**	Malawi
BB	Barbados	**HU**	Hungary	**NL**	Netherlands
BE	Belgium	**IT**	Italy	**NO**	Norway
BG	Bulgaria	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		
FR	France	**ML**	Mali		

AN IMPROVED HEAT-SHOCK CONTROL METHOD AND SYSTEM FOR THE PRODUCTION OF COMPETENT EUKARYOTIC GENE PRODUCTS.

INTRODUCTION

The use of and the advantages of the heat-shock expression elements for the production of gene-products of interest under heat-shock control in a variety of cell types, have previously been described DREANO et al. gene (1986), in press.

This technique has the intrinsic advantages as a eukaryotic expression system of having an efficient, general and highly inducible character. These factors are of considerable interest for the commercial production of proteins of pharmaceutical and other interest, particularly where these proteins are complex, modified, unstable or potentially toxic to the producer cell.

The natural heat-shock system in the living cell performs, amongst other roles, a complex and as yet poorly defined protective function; the nature of heat-shock transcriptional and translational heat-shock control elements have presumably evolved in a fashion leading to an optimization of the physiological role of their associated gene products. The characteristics of production and stability of heat-shock messenger RNA's and proteins are specific to their physiological roles, and are thus not necessarily those optimally desired for the heat-shock controlled production of proteins of interest. Since a number of sequence domains involved in heat-shock promotor regions as well as heat-shock RNA leaders have been described see e.g. AMIN et al., Mol. Cell. Biol. 5, (1985), 197-203, it appeared to us that sequence modifications in the non-coding heat-shock control constructions could lead to advantageous variant sequence constructs for the production of proteins of interest. Also, for practical purposes it may also be advantageous to lower the threshold of temperature or of other stresses for high level expression: such a modification may render the induction of genes of interest under heat-shock control more economical in situations of large scale culture. Moreover, producer cells exposed to a lesser degree of stress may be better capable of supporting high levels of protein synthesis without damaging effects possibly arising from stress.

In addition, the heat-shock controlling elements have also been selected from the human genome to improve compatibility between the expression system involved and the selected host-cell machinery.

SUMMARY OF THE INVENTION

The modifications which led to the method of the invention as summarized by claim 1 can be mainly of three kinds to be applied independently or in combination, i.e. nucleotide insertion, nucleotide deletion and nucleotide substitution. Insertion and deletion may include single nucleotides or polynucleotide fragments. The 3' and/or 5' non-translated sequences can be derived from hsp genes, foreign genes or synthetic DNA such as linkers and adapters and involve ribosome binding and translation initiation sites sequences adjacent to the ATG initiator. Some modifications also relate to the hs consensus sequence identified by PELHAM et al, Trends Genet 1 (1985), 31-34, in the hsp 70KDa of Drosophila Melanogaster, of formula CnnGAAnnTTCnnG; this sequence proposition is by no means exclusive, and is the proposal of these authors, based on the data available to them at the time. As our data base of heat-shock consensus sequences expands, we may expect changes in the statistical consensus sequence that can be described. In some embodiments of the invention, the number of such sequences in the 5' non-translated region has been increased over that normally present in the hsp70 KDa gene; in other embodiments, the position of said sequences relative to each other and to other promoter elements has been modified. The "heat shock gene control element" is defined as being composed of both a heat-shock promotor enabling control of DNA transcription into RNA, and a heat-shock RNA leader sequence influencing translational control into protein. It has generally been observed that the activity of a heat-shock promotor is independent of cell type, while the specific heat-shock translational control is largely dependent on the homology of the RNA leader and the cell or organism employed for expression.

The invention provides for a method for enhancing and inducing the expression of a gene, coding for the production of a gene product (comprising a sequence of amino acids, i.e., a protein), under the control of a heat shock gene control element of the type comprising a heat shock gene translational control element, a heat shock gene promoter sequence and leader sequence, said method comprising :

(1) assembling sequences of deoxynucleotides into a fragment insertable into a plasmid and in a 5' to 3' direction comprising, in order, a heat shock gene promoter sequence, a transcription start site sequence, a 5' transcribed but not translated leader sequence for translational control in

protein production, a consensus sequence coding for translation initiation, a gene sequence coding for production of a protein, and a nontranslated sequence at the 3' end,

the heat shock gene promoter sequence comprising not more than 1500 deoxynucleotides and from 2 to 12 heat shock promoter consensus regions, each said heat shock promoter consensus region comprising not more than 11 deoxynucleotides of formula substantially approaching

1	2	3	4	5	6	7	8	9	10	11
C	$\binom{T}{G}$	$\binom{C}{A}$	G	n^1	n^2	n^3	n^4	T	T	C

such that at least 5 deoxynucleotides in the formula in positions 1 to 4, and 9 to 11, are substantially conserved from heat shock promoter consensus region to heat shock promoter consensus region, and wherein each n is independently selected from A, T, C or G,

the consensus sequence coding for translation initiation comprising not more than 20 deoxynucleotides and substantially comprising the formula

-5	-4	-3	-2	-1	1	2	3	4	
n_x	C	C	A	C	C	A	T	G	G	$n_{(y-x)}$

wherein CC and CC predominates in positions -1, -2, and -4, -5 at an occurence frequency of at least 25 %, wherein A predominates in position -3 at an occurence frequency of about 80 %, wherein ATG predominates in positions 1, 2 and 3 at an occurence frequency of about 90 %, wherein G predominates in position 4 at an occurence frequency of 40-60 %,

wherein each n is independently selected from A, T, C, or G, wherein y is selected from 0 to 11, wherein x is selected from 0 to y,

the nontranslated sequence at the 3' end comprising nontranslated gene sequences of not more than 3000 deoxynucleotides, a terminator sequence, and a poly A addition sequence.

(2) inserting the fragment into a plasmid so as to form a heat shock gene promoter controlled plasmid,

(3) inserting the heat shock gene promoter controlled plasmid into a eukaryotic cell competent in the expression of the gene coding for the production of the protein,

(4) stressing the eukaryotic cell to induce expression of the gene coding for the production of the protein.

It is to be understood in step 1 above here and in the claims that the order is important for the DNA sequences assembled in the named order,

however, additional intervening nucleotides can be included between the ordered sequences without detriment.

Preferably the heat shock gene promoter is derived from the heat shock genomic DNA of Drosophilia melangaster, particularly the genes encoding the 70 kilodalton heat shock proteins.

The heat shock gene control element is composed of both a heat shock promoter enabling DNA transcription into RNA and a heat shock leader sequence influencing translational control of RNA into protein. It has generally been observed that the activity of the heat shock promoter is independent of the specific eukaryotic cell, while heat shock translation is dependent on the homology of the RNA leader and the cellular organism used for its expression.

The RNA leader, the 5' transcribed but not translated leader sequence, is between the transcription start site and the translation start site. The leader can be a heat shock leader enabling translation to occur at heat shock temperature or a leader from another source permitting RNA translation at normal growth temperature.

BRIEF DESCRIPTION OF THE FIGURES

The invention will now be described with reference to the drawing in which the symbols defining the DNA sequences are shown, see fig 7.

Fig. 1a represents plasmid p17, a DNA construction derived from pBR322 and incorporating a human heat-shock control sequence isolated from a human λ genomic library (VOELLMY et al., PNAS USA. 82 (1985), 4949-4953).

Fig. 1b represents plasmid pSP65 (New England Nuclear Laboratories) comprising a BamH I restriction site within a polylinker fragment

Fig. 1c represents plasmid p17 65 JO resulting from insertion of the EcoR I (fill-in)-Hind III fragment from p17 into Pvu II-Hind III sites of pSP 65.

Fig. 1d represents plasmid p17 JO resulting from Bgl II digestion and partial digestion with BamHI of p17 65 JO and auto ligation of a resulting 3,4 Kb fragment.

Fig. 2a represents plasmid phGH (obtained from Celltech Ltd., Slough, GB) containing a c-DNA fragment of a human growth hormone (hGH).

Fig. 2b represents plasmid pSP hGH resulting from the insertion of a BamH I fragment of phGH into the BamH I site of pSP65.

Fig. 2c represents plasmid p17 hGH resulting from the insertion of a

Sal I - Hind III fragment of pSP hGH into corresponding restriction sites of the polylinker of p17 JO (fig. 1d).

Fig. 3a illustrates plasmid p17 hGH Δ 6 resulting from the deletion of six nucleotides in the sequence preceding the hGH initiation of translation site.

Fig. 3b represents plasmid pSV2 dhfr (SUBRAMANI et al., Mol. Cell Biol. 1 (1981), 854-864) containing the mouse dihydrofolate reductase gene under the control of the SV40 early promoter sequence and terminated by the SV40 early terminator sequence.

Fig. 3c represents plasmid p17 hGH dhfrΔ 6 obtained by ligation of a 1.2 Kb EcoRI - BamH I fragment of p17 hGH Δ6 into the BamH I and EcoR I sites of pSV2 dhfr.

Fig. 4a represents plasmid pSV2 NEO (disclosed by SOUTHERN et al., J, Mol. Appl. Genet. 1 (1982), 327-341) containing a neomycin (NEO) resistance gene under the control of the early SV40 promoter and terminated by the early SV40 terminator sequence.

Fig. 4b represents plasmid pSP NEO 9 which results from the ligation of a Pvu II - BamH I fragment derived from pSP65 and a fragment resulting from the partial digestion of pSV2 NEO with the same endonucleases (Pvu II - BamH I).

Fig. 4c represents plasmid p17 hGH NEO resulting from the ligation of a 1.3 Kb EcoR I fragment derived from p17hGH (see fig. 2c) with EcoR I linearized pSP NEO 9.

Fig. 5 is an autoradiography of an acrylamide gel electropherogram showing the various protein fractions inducibly expressed in frog oocytes using a selected gene of interest, according to the invention. Lane identification is given below; C means control temperature. Immunoprecipitation is performed with anti-hGH serum.

M	Protein markers		
1	p17 lys	HS	CYTOPLASM
2	p17 hGH	C	"
3	p17 hGH	HS	"
4	p17 hGH Δ 6	C	"
5	p17 hGH Δ 6	HS	"
6	p17 hGH dhfr Δ 6	C	"
7	p17 hGH dhfr Δ 6	HS	"
8	p17 hGH NEO	C	"
9	p17 hGH NEO	HS	"
10	p17 lys	HS	CULTURE MEDIUM
11	p17 hGH	C	"
12	p17 hGH	HS	"
13	p17 hGH Δ 6	C	"
14	p17 hGH Δ 6	HS	"
15	p17 hGH dhfr Δ 6	C	"
16	p17 hGH dhfr Δ 6	HS	"
17	p17 hGH-NEO	C	"
18	p17 hGH-NEO	HS	"
M			Protein markers

Fig. 6 refers to data similar to that illustrated in fig. 5 but genes are expressed in COS (monkey) and CHO (hamster) cells.

1	p17 lys	HS	COS
2	p17 hGH dhfr Δ 6	HS	COS
3	p17 hGH dhfr Δ 6	C	COS
13	p17 lys	HS	CHO
14	p17 hGH dhfr Δ 6	HS	CHO
15	p17 hGH dhfr Δ 6	C	CHO

The position of protein molecular weight markers (in kilodaltons, KDa) are indicated on the right hand side of the figure.

Fig. 7 is a key to DNA sequence symbols.

Fig. 8 illustrates the base sequences of a region of the Drosophila hsp 70KDa gene including the transcriptional and translational site and the 5' non-transcribed sequences including the promoter with the four consensus region sequences (-49 to -67, -64 to -87, -130 to -195, -191 to -430).

The plasmids identified in figs 1a to 4c include the following sequences and restriction sites (see fig. 7)

Plasmid	Sequences : Kb
p17	11 : 2.6 ; 10 : 3.9
pSP65	11 : 3.0 ; 15 < 0.1
p17 65 JO	11 : 2.8 ; 10 : 3.2 ; 15 < 0.1
p17 JO	11 : 2.8 ; 10 : 0.6 ; 15 < 0.1
phGH	11 : ~3 ; 12 : 0.7
pSP hGH	11 : 3.0 ; 12 : 0.7 ; 15 < 0.1
p17 hGH	11 : 2.8 ; 10 : 0,6 ; 12 : 0.7
p17 hGH Δ6	ditto
pSV2 dhfr	11 : 2.3 ; 13 (promoter):0.4 16 : 0.7 ; 13 (terminator):1.6
p17 hGH dhfr	11 : 2.8 ; 13 (promoter): 0.4 16 : 0.7 ; 13 (terminator):0,7 10 : 0.6 ; 12 : 0,7
pSV2 NEO	11 : 2.3 ; 13 (promoter):0.4 14 : 1.5 ; 13 (terminator):1.6
pSP NEO9	11 : 2.8 ; 13 (promoter): 0.4 14 : 1,5 ; 13 (terminator):0.7
p17 hGH NEO	11 : 2.8 ; 13 (promoter): 0.4 12 : 0.7 ; 13 (terminator):0.7 14 : 1.5 ; 10 : 0.6

NI = Nco I
BI = BamH I
XI = Xba I
HIII = Hind III
RI = EcoR I
SI = Sal I
BII = Bgl II
PII = Pvu II

DETAILED DESCRIPTION OF THE INVENTION

In a first embodiment of the invention, a plasmid, p17, derived from pBR322 and containing a functional human hs control element was obtained by screening by plaque hybridization a lambda human genomic library, using as a probe a coding sequence for the Drosophila hsp 70 Kba gene (VOELLMY et al., PNAS USA 82 (1985) 4949-4953). The presence of a functional human hs control element in the DNA isolated in the cloning procedure was determined by sequence analysis as disclosed in EP-A-118 393 and the foregoing VOELLMY et al. (1985) PNAS reference. A restriction fragment containing a 3.15 Kb hs non transcribed sequence and a 115 bp RNA leader was ligated with a Pvu II-Hind III fragment from pSP65 (fig 1b, a commercially available plasmid) to provide plasmid p1765 JO (fig. 1c). A BamH I - Bgl II restriction fragment therefrom (3.4 Kb) containing a small 5' non transcribed sequence (0.45 kb) was autoligated to provide plasmid p17 JO (fig. 1d).

A human growth hormone c-DNA containing clone (plasmid phGH obtained from Celltech Ltd., Slough, GB, see fig. 2a) was inserted into the polylinker of pSP65 (fig. 1b) to provide plasmid pSP hGH (fig. 2b). Then a Hind III - Sal I 0,7 Kb fragment of the latter comprising the hGH sequence was inserted into the corresponding restriction site of the polylinker of p17 JO to give p17 hGH (fig. 2c); in this construction the hGH sequence is separated from the hs sequence by 27 bp including an ATG triplet before the initiation site. By digestion with Xba I and Nco I and filling in with DNA polymerase large fragment, deletion of six nucleotides immediately preceding the ATG triplet was accomplished, which was followed by auto-ligation to provide plasmid p17 hGH Δ 6 (fig. 3a). This plasmid contains the hGH sequence under the control of the human hs promoter, with a RNA leader sequence modified by deletion. This sequence contains one for optimal translation initiation, conforming to the established consensus sequence involving strong eukaryotic translation initiation sites (see for instance M. KOZAK, Nature 308 (1984), 241-246). Indeed, enhancement of the expression of the deletion modified hGH gene in suitable eukaryotic hosts compared to a control involving no deletion was observed. This will be described hereafter.

In a second embodiment, a 1.2 Kb restriction fragment from p17 hGH Δ 6 containing the hGH gene and the hs sequence was inserted into the BamH I and EcoR I sites of plasmid pSV2 dhfr (fig. 3b), the latter being obtained as disclosed by SUBRAMANI et al. Mol, Cell Biol. 1 (1981), 854-864. The re-

sulting plasmid p17 hGH dhfr Δ6 (fig 3c) included the hGH sequence under the control of the human hs promoter (Δ 6 included) and terminated by the SV40 late terminator (2) and, secondly, the dhfr gene under the control of the SV40 early promoter (1) (note the inverted reading mode). In this embodiment, the presence of the 3' late untranslated sequence also enhanced expression under stress as will be seen hereafter.

In a third embodiment, a plasmid, pSP NEO 9, (fig. 4b) was constructed from a BamH I-Pvu II restriction fragment of pSV2 NEO (SOUTHERN et al, J. Mol. Appl. Genet. 1 (1982), 327-341) containing the neomycin-resistance gene under the control of the SV40 early promoter and ended by the SV40 early terminator (including the poly-A addition site), and a corresponding restriction fragment from pSP 65. Digestion of p17 hGH (fig. 2c) with EcoRI provided a 1.3 Kb fragment which was ligated with EcoR I linearized pSP NEO 9 to provide plasmid p17 hGH NEO (fig. 4c). This plasmid contains two transcriptions units: first the NEO gene under the control of the SV40 early promoter and, second, the hGH gene under the control of the hs promoter with an unmodified initiation site but terminated by the SV40 early terminator. The presence of the 3' untranslated sequence also enhanced the expression of the hs controlled hGH gene in suitable host cells under stress.

Expression experiments and results will now be discussed hereunder; further construction embodiments will be disclosed hereafter.

In a first series of expression experiments the transient expression system of injection of DNA constructs into Xenopus oocyte nuclei were used as previously described (VOELLMY and RUNGGER, PNAS USA 79, (1982), 1776-1780) to analyse the controlled synthesis of human growth hormone (hGH). The plasmid constructions resulting from the aforementioned embodiments were used. In brief, the injected DNA-containing oocytes were subjected to stress (90 minutes at 36°C followed by incubation at 21°C for 12 hours in the presence of ^{35}S-labelled methionine (500 µCi/ml); the presence of labelled protein products in either the cytoplasmic extract of the oocytes, or in the suspension media, was determined by immunoprecipitation essentially as previously described (EP-0.118.393); the results obtained are shown in fig. 5.

In addition to the protein molecular weight markers indicated in lanes M. two series of analyses of immunoprecipitated proteins are shown; lanes 1-9 indicate the results obtained from the cytoplasmic extracts of the injected oocytes, while the lanes 10-18 show the corresponding results

obtained from the suspension media of the same oocytes. The results obtained concerning hGH synthesis are essentially identical for the two series of analyses, and the results on the secreted, extracellular proteins will be further described in the experimental part.

Lanes 11 and 12 show the results using plasmid p17 hGH at the control and heat-shock temperatures respectively; lanes 13 and 14, 15 and 16, 17 and 18 are comparable experiments employing plasmids p17 hGH Δ6; p17 hGH dhfr Δ6; and p17 hGH NEO. The results may be summarized with respect to the production of a protein of molecular weight ≃ 21 kDa specifically immunoprecipitated by the anti hGH antisera, as follows:

The plasmid p17 hGH, which is an RNA leader fusion construct within the RNA leader sequences derived from the human hsp70 clone and lacking five nucleotides of the human sequence, produces a detectable level of hGH after heat-shock, with no apparent synthesis in the absence of heat-shock. Plasmid p17 hGH Δ 6 is a 5'untranslated sequence modification that provides a theoretically stronger translation initiation site; this plasmid is seen to behave in a manner comparable to the previous one in the transient expression assay employed. It should be remembered that both of these plasmid constructions lack a classical 3' end terminator and poly A addition signal. Plasmid p17 hGH NEO incorporates such a 3' modification by the insertion of the neomycin-resistance gene in an inverted mode with respect to the hGH sequences, and this with the unmodified initiation sequence; here a certain degree of increased efficiency of synthesis is observed. When the two types of 3' and 5' modifications are combined in plasmid construction p17 hGH dhfr Δ6, a dramatic increase in expression efficiency of hGH is observed, and this being almost totally under heat-shock control (see lanes 15 and 16, fig. 5).

Plasmid p17 hGH dhfr Δ 6 was also used to transfect cultures of COS-1 monkey cells (Gluzman, Cell 23 (1981), 175-182). The expression of human hGH in these cells was analysed using procedures essentially as described (DREANO et al., gene (1986) in press). In this case an induction of 3 hours at 43°C followed by incubation at 37°C for a further 16 hours in the presence of ^{35}S-methionine was used. Control cultures that had not been heat-stressed were treated in a parallel fashion; both sets of cell cultures were then processed by immunoprecipitation with anti hGH sera and analyzed on acrylamide gels as described above. The results obtained from analyses of the suspension cultures are shown in fig. 6.

The results obtained from the transfection of COS cells are shown in

lanes 1, 2 and 3, and it can be seen that neither transfection with control DNA followed by a heat-shock (lane 1), or transfection with the optimal hGH construct in the absence of a heat-shock (lane 3), give rise to the production and secretion of a ≃ 21 KDa protein specifically precipitated by anti-hGH sera. The COS cells transfected with p17 hGH dhfr Δ 6 and submitted to a heat-shock produce a single, major secreted protein band of the expected molecular weight for hGH (see lane 2). An essentially identical result is obtained from the transfection studies using CHO cells (chinese hamster ovary), see fig. 6, lanes 13, 14 and 15.

It is therefore demonstrated here that in the technique of directing the synthesis of a heterologous human protein under heat-shock control, relatively limited modifications of the non-translated sequences employed in the expression constructs can lead to increased levels of production of the gene product of interest. It is evident that the sequence variants that have been employed are examples of those that can be produced using analogous and related techniques. The precise level of action of the new sequence elements has not been determined although, in general, modifications of heat-shock related sequences could be expected to act at the levels of promotor activity, RNA stability, and translational efficiency. Clearly other sequence variants can be produced and may have similar or complementary effects on the optimization of heat-shock control constructs for the production of proteins of interest. It has been demonstrated that the human elements function in an analogous way for the controlled synthesis of a non-heat-shock protein to the elements derived from Drosophila (EP-0.118.393). In addition, it is demonstrated that the production of a secretable protein, for example human growth hormone, is possible under human heat-shock control, and that this occurs, together with secretion of the protein out of the cell, and this when the same plasmid construction is introduced into frog, monkey cells and others. It may be assumed that the same or similar constructions will permit the production of many proteins including secreted proteins from a large variety of cell types and organisms, including human cells.

The next embodiments concern modifications involving the consensus sequence present in the 5′ non-translated region of 70 KDa hs protein genes from Drosophila. This consensus sequence of 14 bp has the formula Cnn GAAnnTTCnnG and has been suggested (PELHAM, Trends Genet 1 (1985), 31-34) between positions -49 and -62 in the 5′ nontranscribed region of Drosophila heat-shock protein genes, and which is responsible for the heat-regulated

expression of the genes when they are present in high numbers of copies in heterologous cell types (PELHAM Cell 30 (1982), 517 - 528; MIRAULT et al., EMBO J. 1 (1982), 1279 - 1285). This sequence element is being referred to as the hs consensus sequence.

Experiments by several groups (DUDLER and TRAVERS, Cell 38 (1984), 391 - 398; AMIN et al., Mol, Cell. Biol. 5 (1985), 197 - 203; SIMON et al., Cell 40 (1985), 805 - 817) have suggested that two consensus-like sequences may be mainly responsible for the heat-regulated activity of the Drosophila 70 KDa hsp genes when present in low copy numbers and in homologous cells. Furthermore, the presence of multiple consensus-like sequences in the promoters of several hs genes has been noted; the promoters of Drosophila 70 KDa hsp genes contain four such sequences in the first 300 bp of 5' nontranscribed sequence. The four elements reside in areas of the promoters designated here regions 1 through 4. Region 1 extends from -49 to -67, region 2 from -64 to -87, region 3 from -130 to -195 and region 4 from -191 to -430 (see nucleotide sequence of promoter in Fig. 8).

These observations have led to the hypothesis that hs promoter activity and inducibility not only depend on the exact sequences of the hs regulatory elements but also on the number of such sequences and on their positions relative to each other and to other promoter elements. Our findings indeed suggest that both the position and the number of consensus-like sequences have significant effects on promoter function. In those experiments, the test gene, placed under modified hs control sequences of eukaryotic origin was the β -galactosidase gene of bacterial origin.

The constructions to support the above findings involved 5' promoter deletion mutants (prefix D for Deletion, the numbers after the prefix refer to the lengths in bp of promoter segments) of the Drosophila 70 KDa hsp- E. coli β-galactosidase hybrid gene p622c (see AMIN et al., Mol. Cell. Biol. 5 (1985) 197-203 who reported mutant gene D-194 (hybrid gene 622c) as well as D67 (622a) and D50 (622 n).

A first control mutant, IN10, the promoter of which contains aforementioned consensus regions 1 and 2 only was constructed as follows: D87 DNA was digested with BssH II (cuts at -67), ends were filled in by DNA polymerase large fragment, Bgl II linkers (AGATCTAGATCT) were added, the material was double-digested with Hind III/Bgl II and Cla I/Bgl II and then electrophoresed on a 1% low melting agarose gel. A 2.6 Kb Hind III/Bgl II fragment including region 2 of the promoter of a Drosophilia 70 KDa hsp gene and pSVOd vector sequences (MELLON et al., Cell 27, (1981) 279 - 288),

including the bacterial and SV40 origins of replication and the ampicillin resistance gene and a 1.2 Kb Cla I/Bgl II fragment containing region 1 and the RNA leader region of the 70 KDa hsp gene as well as the first one third of the β-galactosidase-coding sequence were isolated. These two fragments were ligated to a similarly purified 4.8 Kb Cla I/Hind III fragment from D87 which included the remainder of the β-galactosidase gene, heat-shock gene 3' trailer sequences and a 0.35 kb vector segment (BamH I to Hind III in pSVOd). Transformation of E.coli and identification of recombinant plasmids were carried out as described previously (see above AMIN et al. reference).

A second mutant, IN20, whose promoter sequence also contains regions 1 and 2 was constructed from IN10, the DNA of which was partially digested with Bgl II (to cut the Bgl II linker between regions 1 and 2), sticky ends were filled in as before, and a SmaI linker (CCCGGG) was added. The sequence modification therefore involved inserting new base pairs between the two concensus sequences of IN10.

A hybrid test gene , R5a, which contains a promoter with one region 1 and two copies of region 2 was constructed as follows:

IN20 DNA was digested with Xho I, ends were filled in, the DNA subsequently digested with Hind III, and a 6.1 Kb fragment containing the hs promoter region, the β-galactosidase gene and 3' trailer sequences was isolated from a low melting agarose gel. This fragment was ligated to a 2.6 Kb Bgl II/Hind III fragment from IN10 which included promoter region 2 and upstream vector sequences.

Then hybrid genes R6a and R7a which contain promoters with one unit of region one and three or four copies of region 2 were made as follows:

R5a DNA was digested with Pst I and partially with Xho I. The resulting 200 and 250 bp long hs promoter fragments were isolated electrophoretically. A 1.7 kb Sma I/Pst I fragment (which included two copies of region 2 and upstream vector sequences) and a 6.9 Kb Pst I fragment (contained part of the hs RNA leader region, the β-galactosidase gene, 3' trailer and vector sequences) from R5a were isolated and ligated to the above promoter fragments.

Further mutants SE1 to SE12 containing promoters with different numbers of synthetic consensus-like sequences were also constructed:

A synthetic linker AGAAGCTTCT (New England Biolabs, custom-synthesized) was preligated and then ligated to Xho I-digested and blunt-ended D50 DNA.

The structures of all plasmids were verified by extensive restriction analysis. Important regions of many constructs were sequenced (using the method of MAXAM and GILBERT, PNAS. USA 74 (1977), 560 - 564). Key promoter sequences showing the number and the position of the consensus like sequences are illustrated hereafter.

Expression experiments were carried out as follows:

In the experiments described here, the different constructs were introduced into Drosophila melanogaster S3 cells by DEAE-dextran-mediated transfection as described previously (LAWSON et al., Mol. Gen. Genet. 198 (1984), 116 - 124). The transient synthesis of the common product of all hybrid genes used, an E. coli-specific β-galactosidase, was measured one day after transfection. A heat-shock, usually at 36.5°C, was applied prior to the preparation of cytoplasmic extracts and measurements of β-galactosidase activities.

The activities of mutants SE1 - 12, which only contain synthetic consensus-like sequence elements in their promoters, were assayed following a 2 hrs heat-shock at 36.5°C or following incubation of the transfected cells for the same length of time at 25 °C the control temperature. It was found that the activities of the hybrid genes are clearly a function of the number of copies of the synthetic consensus-like sequence present in their promoter regions. Significant heat-regulated activity could only be measured with mutants containing promoters with at least 3 sequence elements. Mutant SE7, containing seven sequence elements, is considerably more active in heat-shocked cells than the control gene IN20 which contains a promoter of a 70KDa hsp gene which includes regions 1 and 2.

Similar results were experienced with hybrid test genes R5a, R6a and R7a (see Experimental part) the promoters of which contains two, three or four copies of region 2, respectively. The results indicated that promoters with two or more copies of region 2 are significantly more active than that of the control gene.

The invention is further illustrated by the following experimental details.

CONSTRUCTION OF PLASMID P17

A human hs 70 KDa gene was isolated as follows from a phage lambda human genomic library (LAWN et al., Cell 15 (1978), 1157-1174). The library is screened by plaque hybridization (BENTON and DAVIS, Science 196, (1977)

180-182) using as a probe a 2 Kb Sal I fragment from plasmid Sal 0 (KARCH et al., J. Mol. Biol. 148, (1981), 219-230) which contains the entire coding sequence for the Drosophila hs70 protein; any other isolate of this coding sequence can be used in place of plasmid Sal 0 as long as it contains regions of adequate practical homology to the human hsp70 gene. This cloning procedure is possible due to the high degree of conservation of the protein-coding sequences of the hs genes during evolution from Drosophila to man (VOELLMY et al., J. Biol. Chem.258 (1983) 3516-3522). Positively hybridizing clones are grown up and their DNA subjected to extensive sequence analysis. The presence of functional human hs control elements in the isolated DNAs is determined in a procedure analogous to that described for identification of Drosophila hs control hybrid plasmids in which the Escherichia coli β-Galactosidase gene was placed under hs control (EP 0.118.393). The hs controlled hybrid plasmids were identified in transient expression experiments using microinjection into Xenopus oocyte nuclei (VOELLMY et al., PNAS. USA 79 (1982) 1776-1780). One plasmid identified using such procedures as carrying human hs control sequences is plasmid pl7, fig. 1a, see VOELLMY et al., PNAS USA (1985), 4949-4953).

In the following section a series of plasmid constructs were prepared in which a c-DNA copy of the sequence encoding human growth hormone is placed under the control of a human hsp70 control sequence.

CONSTRUCTION OF pl7 hGH dhfrΔ 6

Plasmid pl7 was digested with EcoR I, filled in with DNA polymerase A Klenow fragment, and further digested with Hind III. A 3.26 Kb fragment containing a human 3.15 Kb non transcribed sequence and an RNA leader sequence of 115 bp was purified. A Pvu II - Hind III fragment from pSP65, fig. 1b (New England Laboratories) was also purified. The two fragments were ligated to give plasmid pl7 65 JO. This latter plasmid was digested with Bgl II and partiallly with BamH I. A 3.4 Kb fragment was purified and "auto" ligated. This new plasmid pl7 JO contains a small 5' non transcribed sequence of 0.5 kb and the above mentioned RNA leader sequence.

A human growth hormone c-DNA clone contained in a BamH I fragment was inserted into the BamH I site of the pSP65's polylinker to give plasmid pSP hGH.

Plasmid pSP hGH was then digested with Hind III and Sac 1 and the hGH sequence was inserted in the same restriction site of the polylinker of pl7 JO to give pl7hGH, (fig. 2c). In this plasmid construction the hGH sequence

was separated from hs sequences by 27 bp, and the sequence before the initiation site was... TC TAG AGG ATCC ATG G...

The p17 hGH was digested with Xba I and Nco I and filled in with DNA polymerase Klenow fragment, allowing a specific deletion of six nucleotides.

The "auto" ligation allowed the construction of plasmid p17 hGH Δ6 which contains the hGH sequence under the control of the human hs promotor with a potentially modified and "strong" initiation site CTAGCATGG. This sequence conforms closely to the established concensus sequence that has been described for strong eukaryotic translation initiation sites. See the following references for details:

M. KOZAK Nucl Acid Res, 9 (1981) 5232-5252
M. KOZAK Nucl Acid Res, 12 (1984) 857-872
M. KOZAK Nature, 308 (1984) 241-246

Finally p17 hGHΔ 6 was digested with EcoR I and BamH I. The 1.2 Kb fragment containing hGH and hs sequence was inserted by ligation into the BamH I and EcoR I sites of pSV2-dhfr (SUBRAMANI et al., Mol. Cell Biol. 1 (1981) 854-864).

The resulting plasmid p17hGH dhfrΔ 6, included firstly the hGH sequence under the control of the human hs promotor, and terminated by the SV40 late terminator sequence and, secondly, the dhfr gene under the control of the SV40 early promotor.

Construction of plasmid p17hGH NEO

Plasmid pSV2-NEO, 10 µg, (SOUTHERN & BERG, J. Mol. Appl. Genet, 1 (1982), 237-341) was digested with BamH I and partially with Pvu II. The fragment containing the neomycin-resistance gene under the control of the early SV40 promotor and ended by the early SV40 terminator (including the poly A-addition site) was ligated with a Pvu II - BamH I fragment derived from pSP 65, to give plasmid pSP NEO 9.

Plasmid p17 hGH, 5 µg, was digested with EcoR I. The resulting fragment of 1.3 Kb was purified and ligated with pSP NEO 9 "linearized" by EcoR I and treated with phosphatase.

The resulting plasmid, p17 hGH NEO contains two transcription units, first the neomycin-resistance gene (as above) and secondly, the hGH gene under the control of the human hs promoter and ended by the late SV40

terminator (including the poly A-addition site). The initiation sequence is unmodified in this plasmid. The scheme used for these plasmid constructions are indicated graphically in figures 1a to 4c.

Expression Experiments

In this first series of expression experiments we have used the transient expression system of injection of DNA constructs into Xenopus oocyte nuclei as previously described (VOELLMY and RUNGGER, P.N.A.S. USA 79 (1982) 1776-1780) to analyse the controlled synthesis of human growth hormone, using the plasmid constructions described in the above specifications. In brief, 2-5 ng of DNA in a volume of 10 nl was used for the injection of 10-20 Xenopus oocytes. Ten minutes after microinjection, the oocytes were subjected to a heat-shock of 90 minutes at 36°C followed by incubation at 21°C for 12 hours in the presence of ^{35}S-labelled methionine (500 µCi/ml; 200 µl for 10-20 oocytes). The presence of labelled protein products in either the cytoplasmic extract of the oocytes, or in the suspension media was determined by immunoprecipitation of the extracts using rabbit anti hGH antisera (obtained from Dako Corporation, Santa Barbara, California; cat. no A570). The detailed procedure for immunoprecipitation and acrylamide gel electrophoretic analysis are essentially as previously described (EP-0.118.393), the results obtained are shown in fig. 5.

In addition to the protein molecular weight markers indicated in lanes M, two series of analyses of immunoprecipitated proteins are shown; lanes 1-9 indicate the results obtained from the cytoplasmic extracts of the injected oocytes, while the lanes 10-18 show the corresponding results obtained by immunoprecipitation of the suspension media of the same oocytes. The results obtained concerning hGH synthesis are essentially identical for the two series of analyses, and the results of the secreted, extracellular proteins will be described in detail here.

Lane 10, using a comparable construction to p17 hGH, but placing chicken lyosozyme gene sequence (KRIEG et al., J. Mol. Biol. 180, (1984) 615-643) under human hs control, no specific protein bands after precipitation with the anti hGH sera, even after the cells have been subjected to a heat shock as described above. Lanes 11 and 12 show the results using plasmid p17 hGH at the control and heat-shock temperatures respectively; lanes 13 and 14, 15 and 16, 17 and 18 are comparable experiments employing plasmids p17 hGHΔ 6; p17 hGH dhfrΔ 6; and p17 hGH NEO. The results may be summarized with respect to the production of a protein of molecular weight

≃ 21 KDa, specifically immunoprecipitated by the anti hGH sera, as follows:

The plasmid p17 hGH, which is an RNA leader fusion construct within the RNA leader sequences derived from the human hsp70 clone and lacking five nucleotides of the human sequence, produces a detectable level of hGH after heat-shock, with no apparent synthesis in the absence of heat-shock. Plasmid p17 hGHΔ 6 is a 5′ untranslated sequence modification that provides a theoretically stronger ribosome binding site; this plasmid is seen to behave in a manner comparable to the previous one in the transient expression assay employed. It should be remembered that both of these plasmid constructions lack a classical 3′ end terminator and poly A addition signal. Plasmid p17 hGH NEO incorporates such a 3′ modification by the insertion of the neomycin-resistance gene in an inverted mode with respect to the hGH sequences, and this with the unmodified initiation sequence; here a certain degree of increased efficiency of synthesis is observed. When the two types of 3′ and 5′ modifications are combined in plasmid construction p17 hGH dhfr Δ 6, a dramatic increase in expression efficiency of hGH is observed, and this being almost totally under heat-shock control (see lanes 15 and 16, fig. 5).

As a demonstration of the generality of the modified hs control construct plasmid p17 hGH dhfr Δ6 was used to transfect cultures of COS-1 monkey cells (GLUZMAN, Cell 23, (1981) 175-182) using the $CaCl_2$-DNA precipitation procedure (GRAHAM and VAN DER EBB, J. Virol., 52 (1973) 456-467). The expression of hGH in these cells was analysed using procedures essentially as described elsewhere (EP-0.118.393). For the transfection of Chinese hamster ovary (CHO) cells, an analogous procedure was employed but the period of amplification used for the transfected plasmids in the COS cells system was omitted. Sixteen hours after transfection for the CHO cells, and fourty hours after for the COS cells, part of the cultures were subjected to a heat-shock of 3 hours at 43°C followed by incubation at 37°C for a further 16 hours in the presence of ^{35}S-methionine (100 µCi/ml; 2 ml/10^7 cells). Control cutures that had not been heat-shocked were treated in a parallel fashion; both sets of cell cultures were then processed in the following way.

Briefly, cytoplasmic extracts were prepared using a treatment comprising the suspension of cells in NET buffer (EP-0.118.393) followed by gentle pipetting on ice and low speed centrifugation; cell suspension cultures were simply clarified by a similar centrifugation step. Samples were then immunoprecipitated with anti hGH sera and analysed on acrylamide

gels as described above. The results obtained from analyses of the suspension cultures are shown in fig. 6.

The results obtained from the transfection of COS cells are shown in lanes 1, 2 and 3, and it can be seen that neither transfection with the control lysozyme construction followed by a heat-shock (lane 1), or transfection with the optimal hGH construct in the absence of a heat-shock (lane 3), give rise to the production and secretion of a specifically anti-hGH precipitated protein. The COS cells transfected with p17 hGH dhfr Δ6 and submitted to a heat-shock produce a single, major secreted protein band of the expected molecular weight (see lane 2). An essentially identical result is obtained from the transfection studies using CHO cells, see figure 6, lanes 13, 14 and 15.

Preparation of mutants IN10, IN20, SE1 to SE12, R5a to R7a and R5b to R7b.

All the above plasmids were prepared according to the schemes disclosed heretofore employing usual means. The sequences of bases of some of the above mutant promoters regions modified according to the invention are listed below; in the formulae, the stars indicate the placement of the consensus sequences; The synthetic sequences were AGAAGCTTCT.

```
                          ****     ***                    ****     ***
IN10          cctcgaggCTGCTCTCGTTGGTTCCAGAGAGCGCGAGATCTCGCGCCTCGAATGTTCGCGA
(-48)

IN20
              ****       ***                              ****       ***
cctcgaggCTGCTCTCGTTGGTTCCAGAGAGCGCGAGATCCCCGGGGATCTCGCGCCTCGAATGTTCGCGA

R5a
              ****     ***                        ****     ***
cctcgaggCTGCTCTCGTTGGTTCCAGAGAGCGCGAGATCtcgaggCTGCTCTCGTTGGTTCCAGAGAGCGCGAG
****     ***
ATCCCCGGGGATCTCGCGCCTCGAATGTTCGCGA
```

R5b

```
          ****     ***                                    ****     ***
cctcgaggCTGCTCTCGTTGGTTCCAGAGAGCGCGAGATCCCCtcgaggCTGCTCTCGTTGGTTCCAGAGAGCGC
****     ***
GAGATCCCCGGGGATCTCGCGCCTCGAATGTTCGCGA
```

R6a

```
              ****     ***                                ****     ***
cctcgaggCTGCTCTCGTTGGTTCCAGAGAGCGCGAGATCtcgaggCTGCTCTCGTTGGTTCCAGAGAGCGCGAG
****     ***                                    ****     ***
ATCCCCtcgaggCTGCTCTCGTTGGTTCCAGAGAGCGCGAGATCCCCGGGGATCTCGCGCCTCGAATGTTCGCGA
```

R7a

```
              ****     ***                                ****     ***
cctcgaggCTGCTCTCGTTGGTTCCAGAGAGCGCGAGATCtcgaggCTGCTCTCGTTGGTTCCAGAGAGCGCGAG
****     ***                                ****     ***
ATCCCCtcgaggCTGCTCTCGTTGGTTCCAGAGAGCGCGAGATCtcgaggCTGCTCTCGTTGGTTCCAGAGAGC
****     ***
GCGAGATCCCCGGGGATCTCGCGCCTCGAATGTTCGCGA
```

SE7

```
cctcgaAGAAGCTTCTAGAAGCTTCTAGAAGCTTCTAGAAGCTTCTAGAAGCTTCTAGAAGCTTC
               ****     ******     ******     ******     ******     ***
TAGAAGCTTCTtcgaggCGA
***     ***
```

Expression experiments:

As already mentioned before, the different construct involving the consensus sequence and copies thereof were tested in Drosophila S3 cells after transfection by the DEAE-dextran technique of LAWSON et al., Mol. Gen. Genet. 198 (1984), 116-124.

The activities of mutants SE1 - 12, which only contain synthetic consensus-like sequence elements in their promoters, were assayed following a 2 hrs heat-shock at 36.5 °C or following incubation of the transfected cells for the same length of time at 25 °C, the control temperature. The results of these experiments are presented in Table 1.

TABLE 1

Mutants	Nb. of sequence elements	B - Galactosidase (relative units)	
		Heat induced	Control
IN20	2 (control)	1.00	0.00
SE1	1	0.05	0.05
SE2	2	0.10	0.00
SE3	3	0.45	0.00
SE7	7	1.80	0.05
SE9	9	0.65	0.05
SE12	12	0.80	0.00

The activities of the hybrid genes are clearly a function of the number of copies of the synthetic consensus-like sequence present in their promoter regions. Significant heat-regulated activity could only be measured with mutants containing promoters with at least 3 sequence elements. Mutant SE7, containing seven sequence elements, is considerably more active in heat-shocked cells than the control gene IN20 (this hybrid gene contains a promoter of a 70 KDa hsp gene which includes regions 1 and 2).

Similar experiments were carried out with heat-shock hybrid genes IN20, R5a, R6a and R7a the promoters of which contain one, two, three or four copies of region 2, respectively. Drosophila cells that had been transfected with the hybrid genes were either exposed to a heat-shock at 36.5°C for 90 or 120 min or were incubated further at 25°C. To exclude the possibility that differences in the activities of the promoters would be obscured by slower post-transcriptional event, transcription was stopped (with actinomycin D) after 15 to 45 min of heat-shock while translation was allowed to continue to the end of the heat treatment. The results of these experiments (Table 2) strongly suggest that promoters with two or more

copies of region 2 are three to five times more active than that of the control gene (IN20) which only contains one copy of this region. Again, it was observed that all constructs are perfectly heat-inducible: they are essentially inactive at the control temperature of 25°C.

However, at 34°C constructs R5a and R6a were found to be expressed at elevated levels. It should be mentioned that IN20 activity is only barely detectable at 34°C. Thus modifications in heat-shock promotor sequences can lower the threshhold of stress induction, i.e. induction temperature in this case, while remaining totally inactive at the control temperature.

These results indicate a minimal heat-shock regulatory sequence as CTCGnnnnTTC. Comparison to Drosophila hsp 84 promotor sequences suggests that the inverted repeat of CTCG, i.e. CGAG is also an acceptable part of heat-shock consensus elements (Holmgren et al., PNAS USA. 78 (1981) 3 775-3778).

TABLE 2

Treatments of cells	β - Galactosidase activity (rel. units)			
	IN20	R6a	R7a	R5a
Incubation at 25°C	0.00	0.05	0.05	-
Incubation at 34°C for 2 hours	0.20	1.50	-	-
Incubation at 36.5°C for 2 hours, actinomycin D added after 15 min of incubation at high temp.	0.30	1.20	1.40	-
ditto, actinomycin added after 30 min	1.00	4.00	3.35	-
Incubation at 25°C	0.00	-	-	0.00
Incubation at 34°C for 2 hours	0.20	-	-	1.05
Incubation at 36.5°C for 90 min, actinomycin D added after 45 min	1.00	-	-	3.55

These experiments demonstrate that significant promoter activity requires the presence of multiple consensus-like sequences, and that promoters with higher activities than the natural ones can be constructed. It is important to note that heat-inducibility is maintained in the synthetic promoters irrespective of the number of consensus-like sequences.

The experiments described here show that the performance of natural heat-shock gene promoters can be improved by <u>in vitro</u> sequence manipulation

of these promoters. It has been demonstrated here that the maximal activity as well as the activity at suboptimal temperatures of a Drosophila heat-shock consensus-like sequences are located near the start of transcription site of the heat-shock gene.

Interestingly, it appears that the addition of a third consensus-like sequence to a promoter of a Drosophila 70 KDa hsp gene including regions 1 and 2 has a more significant effect than the subsequent addition of a forth or fifth element (Table 2).

Although we have not systematically tested effects of the distance between the newly added heat-shock elements and the downstream promoter elements on promoter activity, the experiments presented here suggest the importance of this parameter. Addition of a third heat-shock consensus sequence in R5a increases promoter activity substantially (see results in Table 2). A third element is located about 100 bp upstream from the second element in the natural Drosophila promoter (see region 3 in Fig. 1). Removal of this third element has been reported not to affect significantly the activity of the promoter (AMIN et al., Mol. cell. biol 5 (1985) 197-203). It thus appears that the natural third element may be too far away from downstream promoter sequences to play an important role in promoter function.

Given the evolutionary conservation of the heat- shock response, it can be assumed that similar manipulations will also improve the performance of heat-shock promoters isolated from other eukaryotic organisms in eukaryotic cells (from Drosophila, man, or other eukaryotic organisms).

Combinations of sequence variants comprising heat-shock promotor, RNA leader, translation initiation and 3'- non-translated sequences can be assumed under certain combinations to provide optimal heat-shock expression levels of genes of interest.

Plasmids of interest have been deposited at the National Collections of Industrial & Marine Bacteria Ltd.

Aberdeen, Scotland, as follows :

Microorganism	Strain number	NCIB number
Escherichia coli	HB101 p17 hGH.NEO	12122
Escherichia coli	HB101 p17 hGH	12123
Escherichia coli	HB101 p17 hGH dhfr Δ6	12124
Escherichia coli	HB101 p17 hGH Δ6	12125

CLAIMS

1. A method for controllably inducing the expression of eukaryotic genes of interest in suitable eukaryotic host cells, this method comprising the steps of:

a) Providing a recombinant DNA gene expression unit in which said gene of interest is functionally linked under the transcriptional and/or translational control of at least one heat-shock control element of eukaryotic origin and including one or more non-translated sequences 5' and/or 3' to said gene of interest;

b) Introducing one or more sequence variance or modification into said non-translated sequences for enhancing the efficiency of said expression upon induction;

c) Introducing said modified gene expression unit into suitable host cells or eukaryotic origin and subjecting the culture to stress whereby the expression of said genes into products of interest is induced.

2. The method of claim 1, wherein the 5' sequences are derived from hsp, foreing genes or synthetic DNA such as linkers and adapters.

3. The method of claim 2, wherein said sequence variance or modification involves site and/or sequences adjacent to the ATG initiator.

4. The method of claim 3, wherein said sequence variance or modification comprises nucleotide insertion /deletion and/or substitution of synthetic DNA.

5. The method of claim 1, wherein the 3' non translated sequences are from terminator regions of hsp or foreign genes.

6. The method of claim 2, wherein the 5' sequences belong to the consensus sequences regulating the expression of the hsp genes of formula approximating CnnGAAnnTTCnnG.

7. The method of claim 6, which comprises increasing the number of the consensus sequence in the promoter region over that number originally present and/or modifying the length of the sequences separating the consensus sequences.

8. A "gene expression unit-host cell combination system" wherein the "gene expression unit" is a DNA construction which comprises.

a) a gene of interest

b) at least one heat-shock protein gene control element of eukaryotic origin for controlling the expression of said gene of interest;

c) 5' and/or 3' end non-translated sequence including modifications

improving the efficiency of expression of genes of interest;

d) at least one site with which the gene of interest is associated to come under the transcriptional and/or translational control of said expression unit, and wherein the "host cell system" comports cells or organisms which are of eukaryotic origin and in which the said expression control unit, functionally linked to the said gene of interest, is present in order to produce the protein product of the said gene.

9. A "gene expression unit-host cell combination" as defined in claim 8, wherein the gene expression unit is derived from human genomic DNA.

10. A "gene expression unit-host cell combination" as defined in claim 8, wherein the gene expression unit is derived from the 70KDa hs genes of eukaryotic cells or organisms.

11. A "gene expression unit-host cell combination" as defined in claim 8 wherein the gene expression unit is derived from the 70Kd hs genes isolated from human genomic DNA.

12. A "gene expression unit-host cell combination" as defined in claim 8 wherein the gene expression unit is associated with a selectable and/or amplifiable genetic marker.

13. Plasmid pl7hGH as a new composition of matter

14. Plasmid pl7hGH Δ6 as a new composition of matter

15. Plasmid pl7hGH dhfr Δ6 as a new composition of matter

16. Plasmid pl7hGH NEO as a new composition of matter.

17. Mutant plasmids SE1, SE2, SE3, SE7, SE9, SE12, R6a, R5a, R7a as new compositions of matter.

18. A method for inducing the expression of a gene, coding for the production of a protein, under the control of a heat shock gene promoter comprising :

(1) assembling sequences of deoxynucleotides into a fragment insertable into a plasmid and in a 5' to 3' direction comprising, in order, a heat shock gene promoter sequence, a transcription start site sequence, a 5' transcribed but not translated leader sequence for translational control in protein production, a consensus sequence coding for translation initiation, a gene sequence coding for production of a protein, and a nontranslated sequence at the 3' end,

the heat shock gene promoter sequence comprising not more than 1500 deoxynucleotides and from 2 to 12 heat shock promoter consensus regions, each said heat shock promoter consensus region comprising not more than 11 deoxynucleotides of formula substantially approaching

1	2	3	4	5	6	7	8	9	10	11
C	$\begin{pmatrix} T \\ G \end{pmatrix}$	$\begin{pmatrix} C \\ A \end{pmatrix}$	G	n^1	n^2	n^3	n^4	T	T	C

such that at least 5 deoxynucleotides in the formula in positions 1 to 4, and 9 to 11, are substantially conserved from heat shock promoter consensus region to heat shock promoter consensus region, and wherein each n is independently selected from A, T, C, or G,

the consensus sequence coding for translation initiation comprising not more than 20 deoxynucleotides and substantially comprising the formula

	-5	-4	-3	-2	-1	1	2	3	4	
n_x	C	C	A	C	C	A	T	G	G	$n_{(y-x)}$

wherein CC and CC predominates in positions -1, -2, and -4, -5 at an occurence frequency of at least 25 %, wherein A predominates in position -3 at an occurence frequency of at about 80 %, wherein ATG predominates in positions 1, 2 and 3 at an occurence frequency of about 90 %, wherein G predominates in position 4 at an occurence frequency of 40-60 %

wherein each n is independently selected from A, T, C, or G, wherein y is selected from 0 to 11, wherein x is selected from 0 to y,

the nontranslated sequence at the 3' end comprising nontranslated gene sequences of not more than 3000 deoxynucleotides, a terminator sequence, and a poly A addition sequence,

(2) inserting the fragment into a plasmid so as to form a heat shock gene promoter controlled plasmid,

(3) inserting the heat shock gene promoter controlled plasmid into a eukaryotic cell competent in the expression of the gene coding for the production of the protein,

(4) stressing the eukaryotic cell to induce expression of the gene coding for the production of the protein.

19. The method according to claim 18 wherein the heat shock gene promoter is derived from the genomic DNA of Drosophila melanqaster.

20. The method according to claim 19 wherein the genomic DNA is the genes encoding the 70 kilodalton heat-shock proteins of Drosophila melanqaster.

21. The method according to claim 18 including in addition in step (1) in the sequences of deoxynucleotides, a sequence coding for a selectable marker for enabling selection of transformant plasmids.

22. The method according to claim 21 wherein the selectable marker is antibiotic resistance.

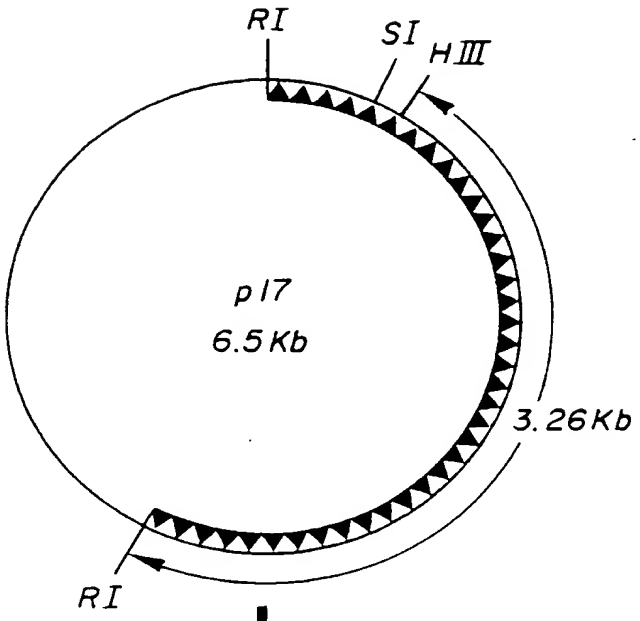


FIG. 1a

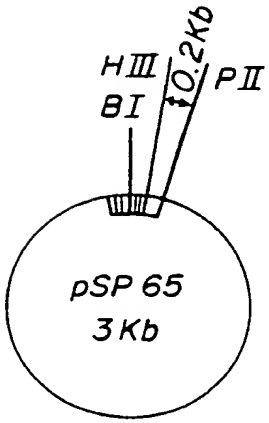


FIG. 1b

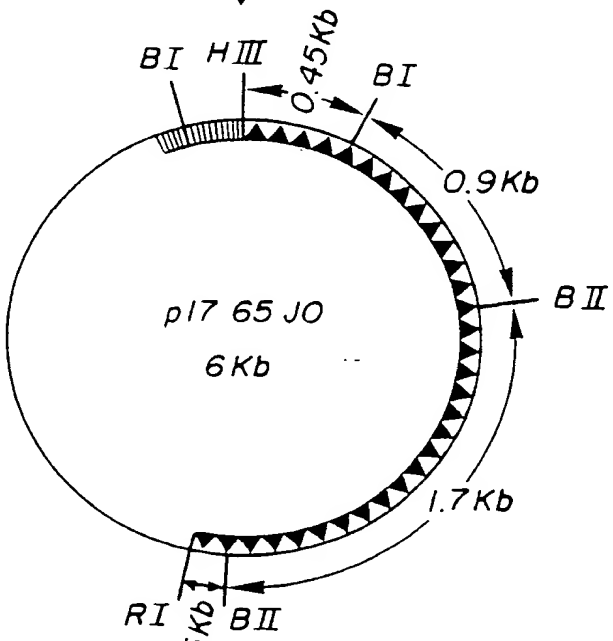


FIG. 1c

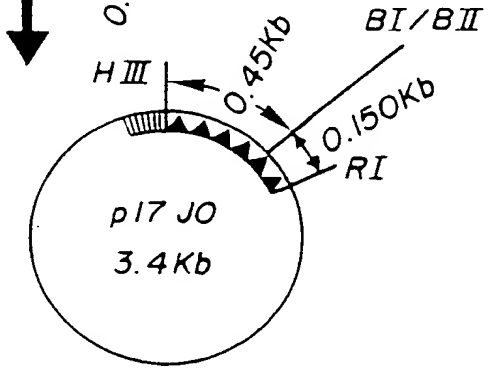


FIG. 1d

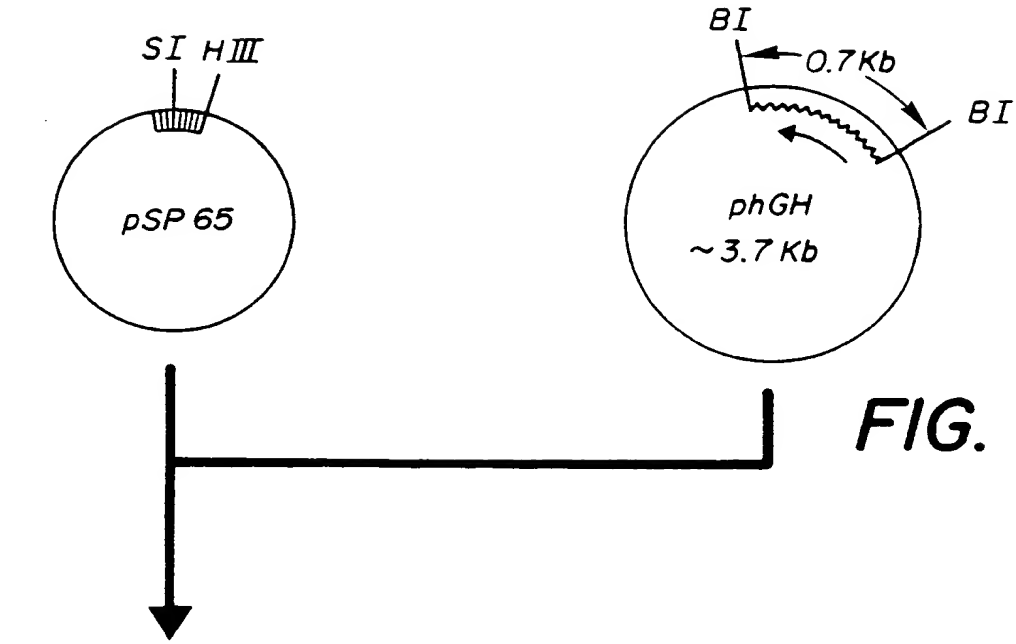


FIG. 2a

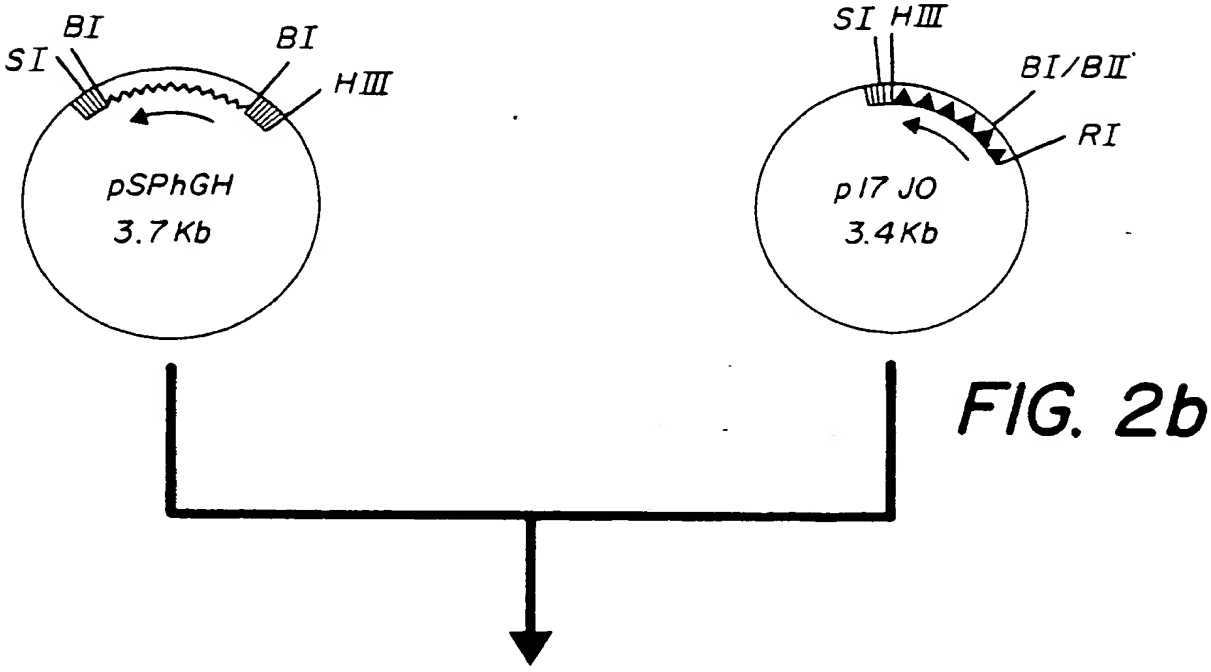


FIG. 2b

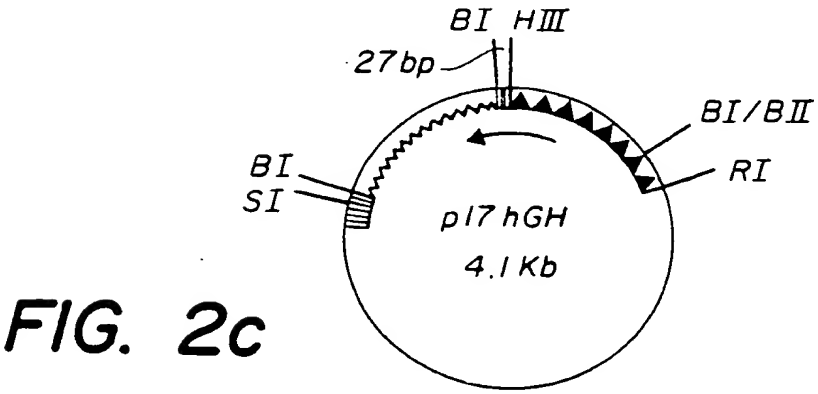


FIG. 2c

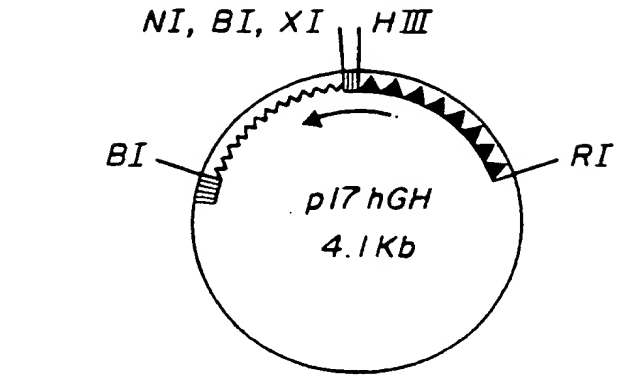


FIG. 3a

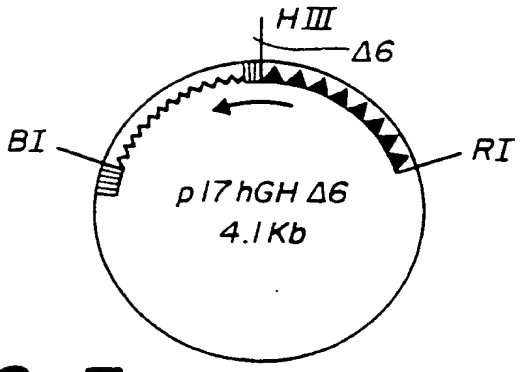


FIG. 3b

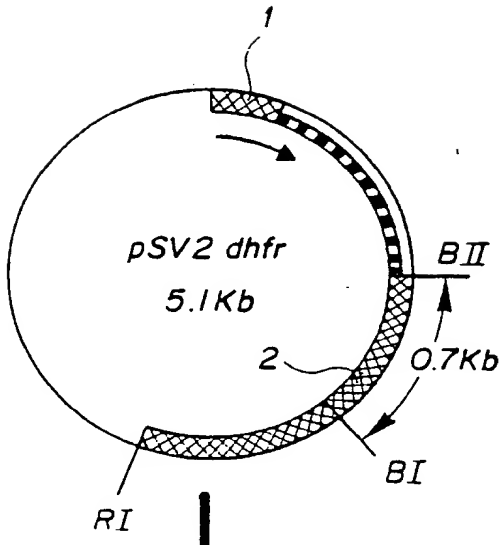
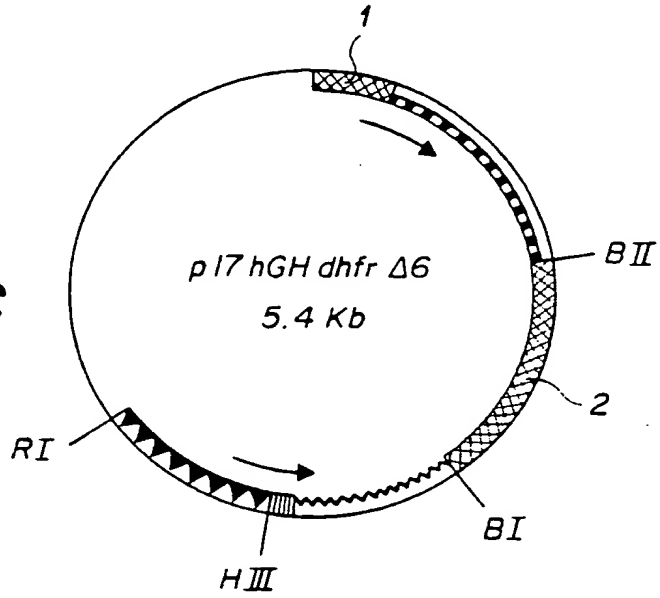


FIG. 3c



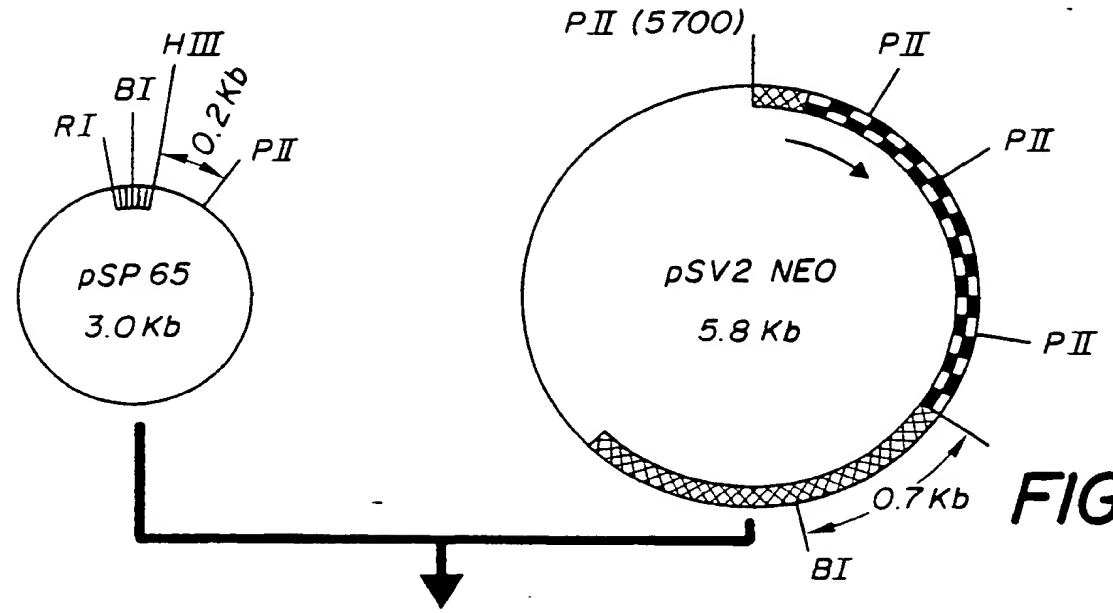


FIG. 4a

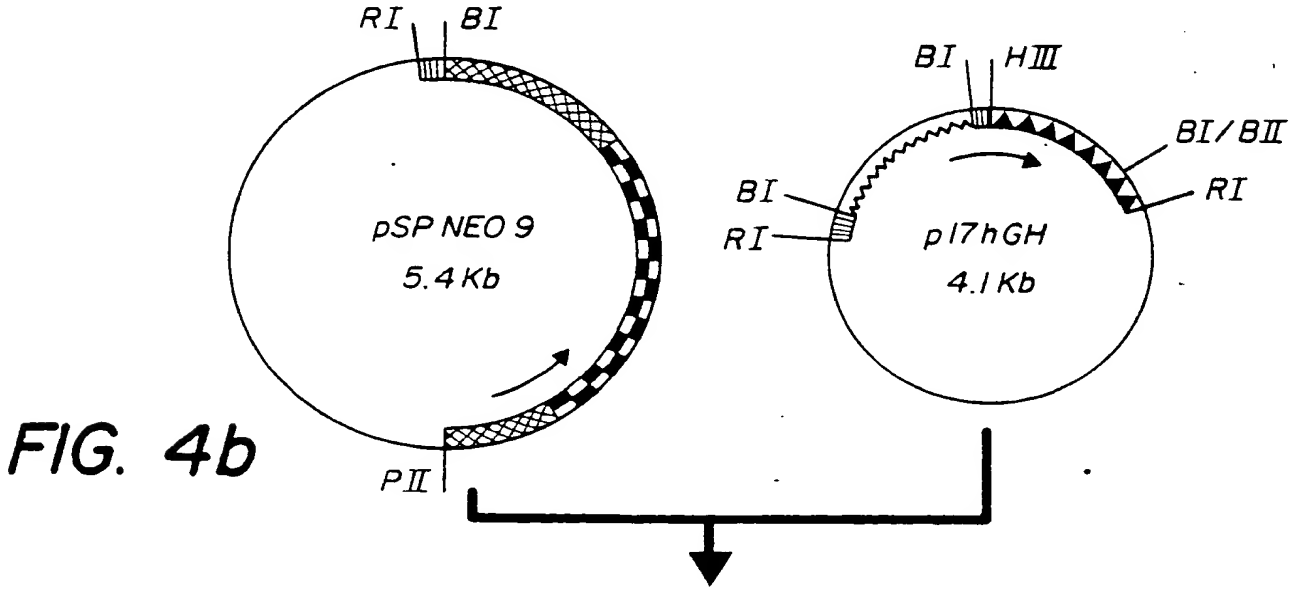


FIG. 4b

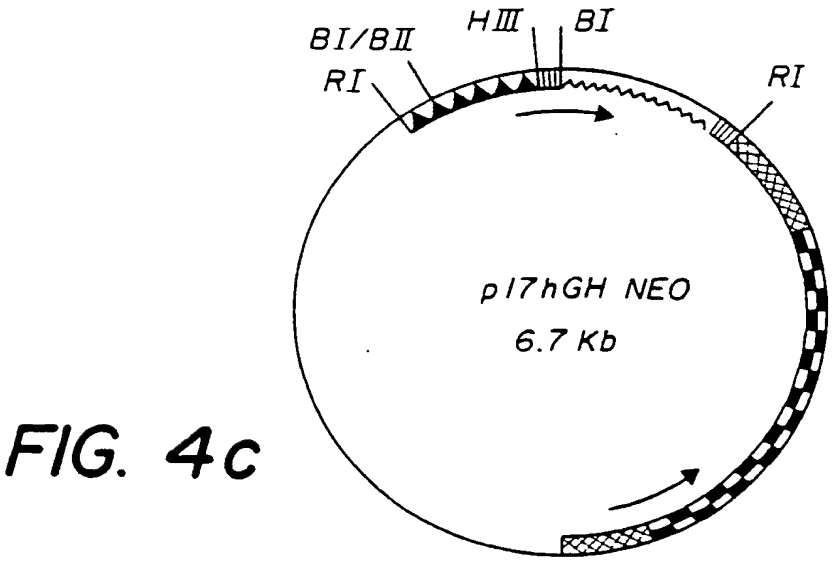


FIG. 4c

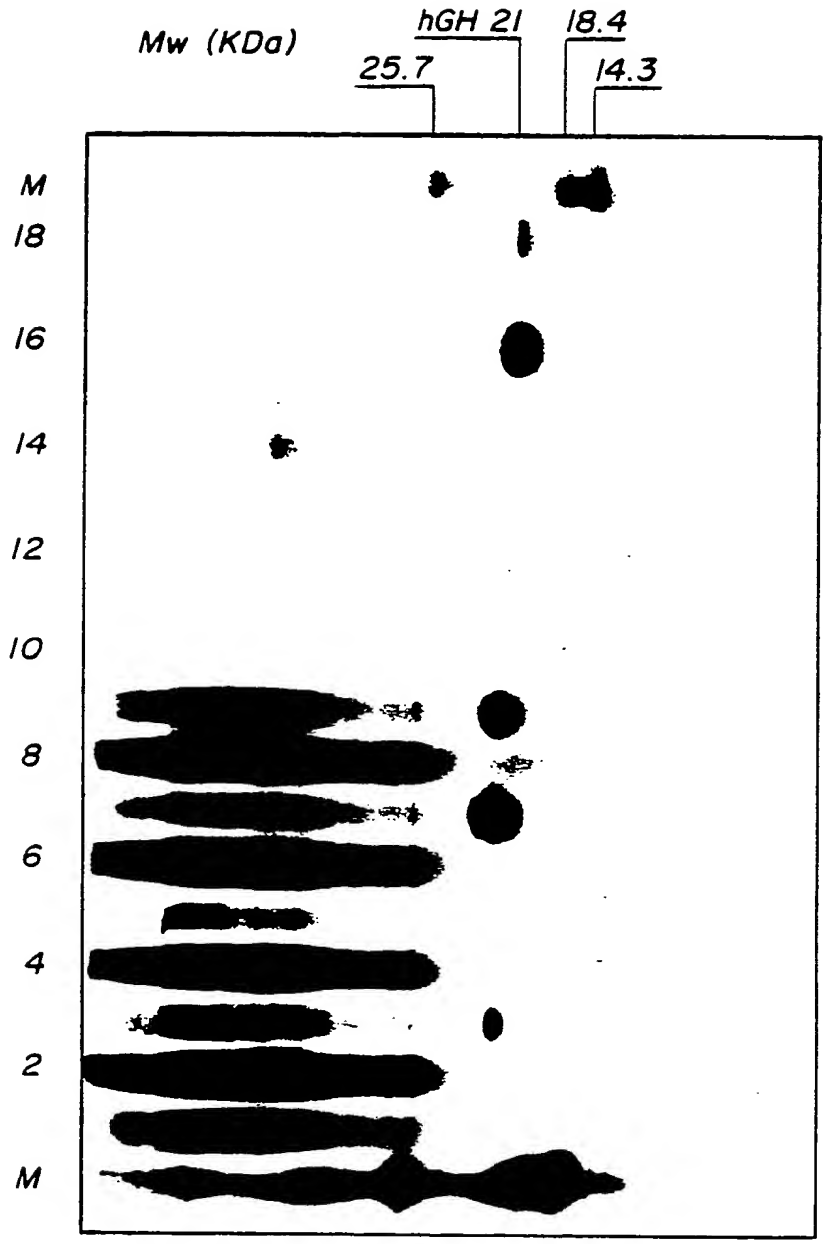


FIG. 5

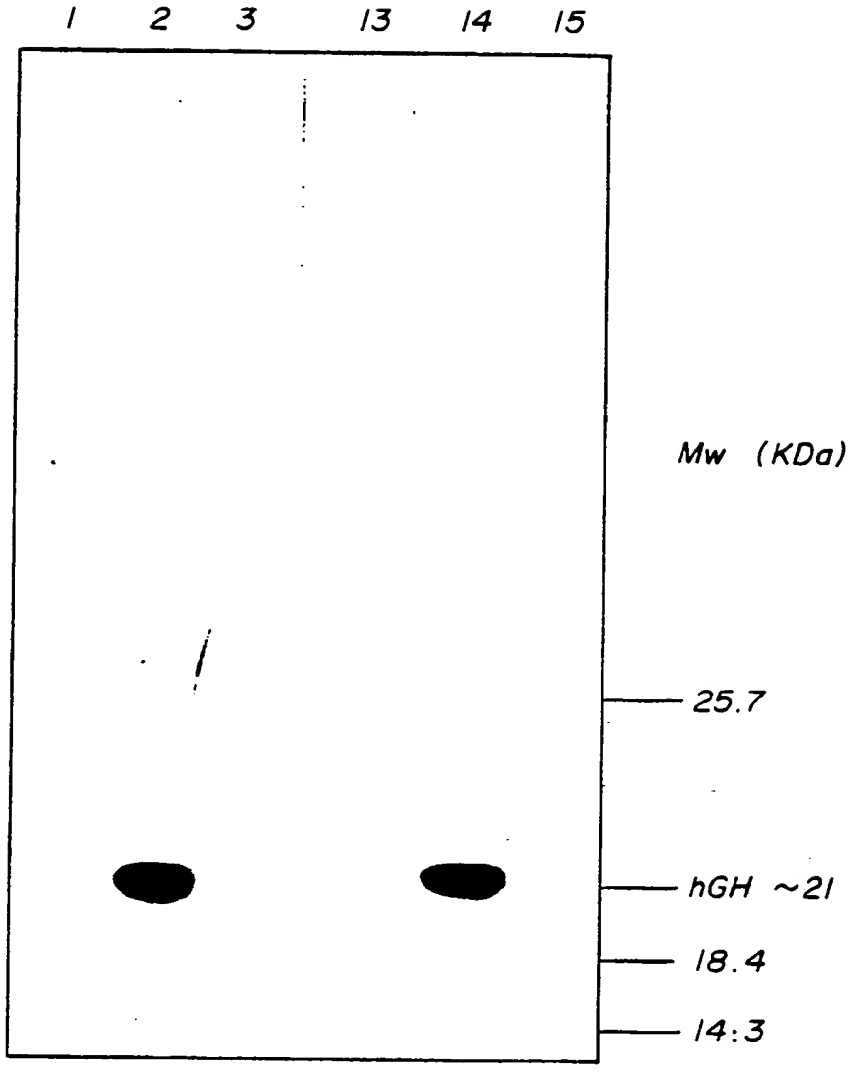


FIG. 6

A KEY TO SEQUENCE SYMBOLS

Human heat shock control sequence

(10)

pBR 322 derived sequence

(11)

Human growth hormone sequence

(12)

SV40 sequence
(promotor and terminator)

(13)

Neomycin gene sequence
(NEO)

(14)

pSP 65 polylinker

(15)

Dihydrofolate reductase gene sequence
(dhfr)

(16)

FIG. 7

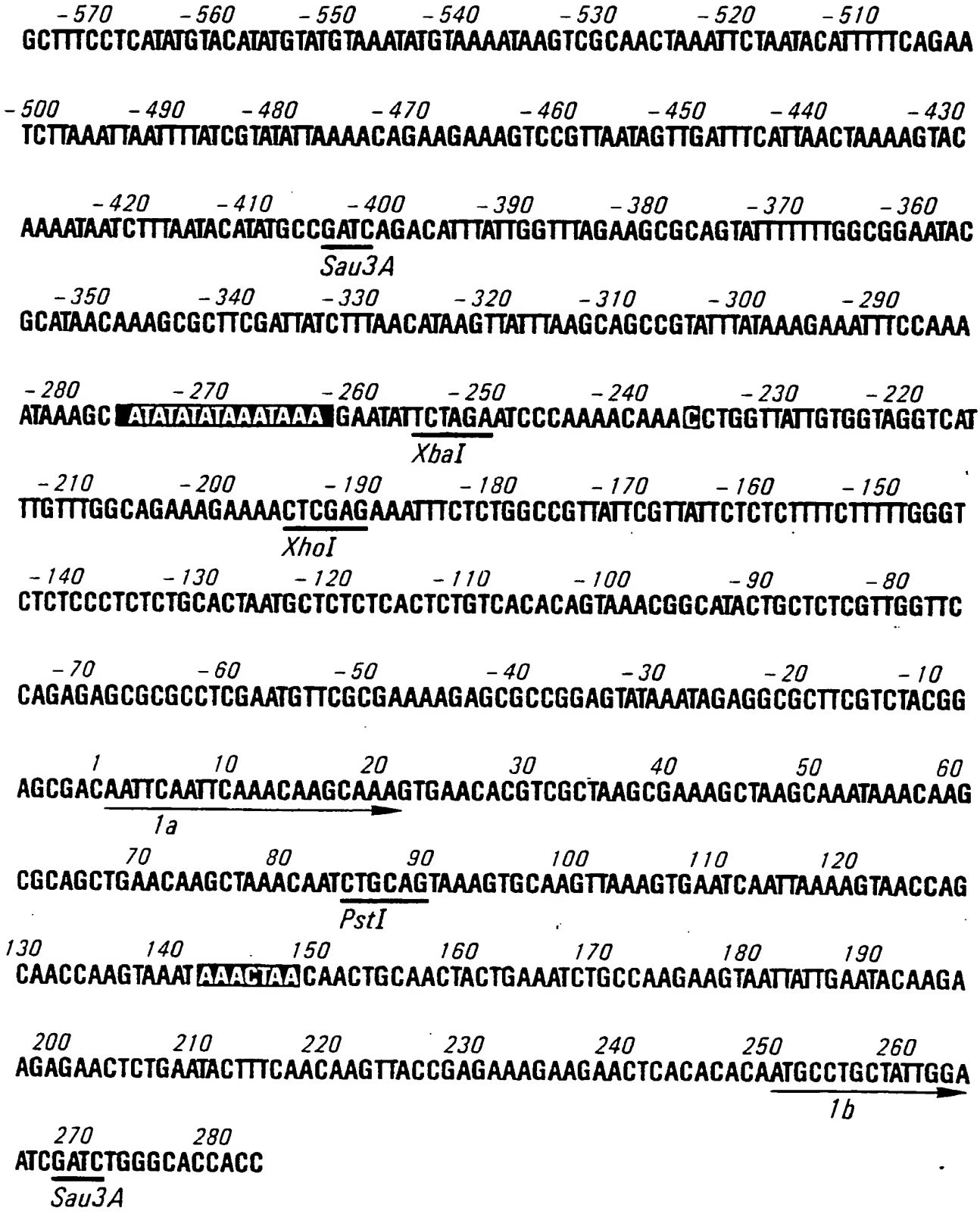


FIG. 8

INTERNATIONAL SEARCH REPORT

International Application No PCT/EP 86/00451

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[4]: C 12 N 15/00

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[4]	C 12 N C 12 P

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [10]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X	The EMBO Journal, volume 1, no. 11, 1982, (IRL Press limited, Oxford, GB), H.R.B. Pelham et al.: "A synthetic heat-shock promoter element confers heat-inducibility on the herpes simplex virus thymidine kinase gene", pages 1473-1477 see the whole article	1-6,8,10,12, 18-21
A	Cell, volume 35, part 1, December 1983, J.T. Lis et al.: "New heat shock puffs and beta-galactosidase activity resulting from transformation of Drosophila with an hsp70-lacZ hybrid gene", pages 403-410 see figure 1	1-12,18-21
A	EP, A, 0118393 (BATELLE MEMORIAL INSTITUTE) 12 September 1984 see the whole document	1-12,18-21
X,P	Proceedings of the National Academy of Sciences USA, volume 82, August 1985, R. Voellmy et al.: ./.	

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
13th November 1986	16 DEC 1986
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	L. ROSSI

Form PCT/ISA/210 (second sheet) (January 1985)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FR M THE SEC ND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
	Isolation and functional analysis of a human 70.000 dalton heat shock protein gene segment", pages 4949-4953 see the whole article (cited in the application)	1-12,18,21

ANNEX TO INTERNATIONAL SEARCH REPORT ON

INTERNATIONAL APPLICATION NO. PCT/EP 86/00451 (SA 14307)

This Annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 21/11/86

The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A- 0118393	12/09/84	JP-A-	59192091	31/10/84
		AU-A-	2420884	17/04/86

For more details about this annex :
see Official Journal of the European Patent Office, No. 12/82